

EXHIBIT A

Joseph R. Saveri (State Bar No. 130064)
JOSEPH SAVERI LAW FIRM, LLP
 601 California Street, Suite 1505
 San Francisco, CA 94108
 Telephone: (415) 500-6800
 Facsimile: (415) 395-9940
 Email: jsaveri@saverilawfirm.com

Matthew Butterick (State Bar No. 250953)
 1920 Hillhurst Avenue, #406
 Los Angeles, CA 90027
 Telephone: (323) 968-2632
 Facsimile: (415) 395-9940
 Email: mb@buttericklaw.com

Laura M. Matson (*pro hac vice* pending)
LOCKRIDGE GRINDAL NAUEN PLLP
 100 Washington Avenue South, Suite 2200
 Minneapolis, MN 55401
 Telephone: (612) 339-6900
 Facsimile: (612) 339-0981
 Email: lmmatson@locklaw.com

*Counsel for Individual and Representative
 Plaintiffs and the Proposed Class
 (continues on signature page)*

**UNITED STATES DISTRICT COURT
 NORTHERN DISTRICT OF CALIFORNIA
 SAN FRANCISCO DIVISION**

Jingna Zhang, an individual;
Sarah Andersen, an individual;
Hope Larson, an individual; and
Jessica Fink, an individual;

Individual and Representative Plaintiffs,

v.

Google LLC, a Delaware limited liability company; and
Alphabet Inc., a Delaware corporation;

Defendants.

Case No.

COMPLAINT

CLASS ACTION

DEMAND FOR JURY TRIAL

Plaintiffs Jingna Zhang, Sarah Andersen, Hope Larson, and Jessica Fink (together “Plaintiffs”), on behalf of themselves and all others similarly situated, bring this class-action complaint (“Complaint”) against defendants Google LLC (“Google”) and Alphabet Inc. (“Alphabet”) (together “Defendants”).

OVERVIEW

1. *Artificial intelligence*—commonly abbreviated “AI”—denotes software that is designed to algorithmically simulate human reasoning or inference, often using statistical methods.

2. Imagen is an AI software product created, maintained, and sold by Google. Imagen is a *text-to-image diffusion model*. A text-to-image diffusion model takes as input a short text description of an image (also known as a *text prompt*) and then uses a machine-learning technique called *diffusion* to generate an image in response to the prompt.

3. Rather than being programmed in the traditional way—that is, by human programmers writing code—a diffusion model is *trained* by copying an enormous quantity of digital images with associated text captions, extracting protected expression from these works, and transforming that protected expression into a large set of numbers called *weights* that are stored within the model. These weights are entirely and uniquely derived from the protected expression in the training dataset. Whenever a diffusion model generates an image in response to a user prompt, it is performing a computation that relies on these stored weights, with the goal of imitating the protected expression ingested from the training dataset.

4. Training a model first requires amassing a huge corpus of data, called a *dataset*. The AI models at issue in this complaint were trained on datasets containing millions of images paired with descriptive captions. In this complaint, each image-caption pair is called a *training image*. During training of the model, the training images in the dataset are directly copied in full and then completely ingested by the model, meaning that protected expression from every training image enters the model. As it copies and ingests billions of training images, the model progressively develops the ability to generate outputs that mimic the protected expression copied from the dataset.

5. Plaintiffs and Class members are visual artists. They own registered copyrights in certain training images that Google has admitted copying to train Imagen. Plaintiffs and Class members never authorized Google to use their copyrighted works as training material.

6. These copyrighted training images were copied multiple times by Google during the training process for Imagen. Because Imagen contains weights that represent a transformation of the protected expression in the training dataset, Imagen is itself an infringing derivative work.

7. Alphabet, as the corporate parent of Google, also commercially benefits from these acts of massive copyright infringement.

JURISDICTION AND VENUE

8. This Court has subject-matter jurisdiction under 28 U.S.C. § 1331 because this case arises under the Copyright Act (17 U.S.C. § 501).

9. Jurisdiction and venue are proper in this judicial district under 28 U.S.C. § 1391(c)(2) because Defendants are headquartered in this district. Google created the Imagen model and, in cooperation with Alphabet, distributes it commercially. Therefore, a substantial part of the events giving rise to the claim occurred in this District. A substantial portion of the affected interstate trade and commerce was carried out in this District. Defendants have transacted business, maintained substantial contacts, and/or committed overt acts in furtherance of the illegal scheme and conspiracy throughout the United States, including in this District. Defendants' conduct has had the intended and foreseeable effect of causing injury to persons residing in, located in, or doing business throughout the United States, including in this District.

10. Under Civil Local Rule 3-2(c), assignment of this case to the San Francisco Division is proper because this case pertains to intellectual-property rights, which under General Order No. 44 is deemed a district-wide case category, and therefore venue is proper in any courthouse in this District.

PLAINTIFFS

11. Plaintiff Jingna Zhang is a photographer who lives in Washington.

12. Plaintiff Sarah Andersen is a cartoonist and illustrator who lives in Oregon.

13. Plaintiff Hope Larson is a cartoonist and illustrator who lives in North Carolina.

14. Plaintiff Jessica Fink is a cartoonist and illustrator who lives in New York.

15. A nonexhaustive list of registered copyrights owned by Plaintiffs is included as **Exhibit A: Plaintiff Copyright Registrations**. A nonexhaustive list of copyrighted images registered by Plaintiffs and infringed by Defendants is included as **Exhibit B: Plaintiff Images in LAION-400M**.

16. The images shown in Exhibit B are offered as a representative sample of works by Plaintiffs that appear in the LAION-400M dataset—not an exhaustive or complete list. Plaintiffs confirmed that these particular images were in the LAION-400M dataset by searching for their own names on two websites that allow searching of the LAION datasets: <https://haveibeenentrained.com> and <https://rom1504.github.io/clip-retrieval/>. On information and belief, all of Plaintiffs' works that were registered as part of the collections in Exhibit A and were online were scraped into the LAION-400M dataset.

DEFENDANTS

17. Defendant Google LLC is a Delaware limited liability company with its principal place of business at 1600 Amphitheatre Parkway, Mountain View CA 94043.

18. Defendant Alphabet Inc. is a Delaware corporation with its principal place of business at 1600 Amphitheatre Parkway, Mountain View CA 94043. In 2015, Google became a subsidiary of Alphabet.

AGENTS AND CO-CONSPIRATORS

19. The unlawful acts alleged against the Defendants in this Complaint were authorized, ordered, or performed by the Defendants' respective officers, agents, employees, representatives, or shareholders while actively engaged in the management, direction, or control of the Defendants' businesses or affairs. The Defendants' agents operated under the explicit and apparent authority of their principals. Each Defendant, and its subsidiaries, affiliates, and agents operated as a single unified entity.

20. Various persons or firms not named as defendants may have participated as co-conspirators in the violations alleged herein and may have performed acts and made statements in furtherance thereof. Each acted as the principal, agent, or joint venture of, or for other Defendants with respect to the acts, violations, and common course of conduct alleged herein.

FACTUAL ALLEGATIONS

21. Google is a diversified technology company whose lines of business include internet advertising and cloud-computing services. As part of these businesses, Google creates and distributes artificial-intelligence software products.

22. One such product is Imagen, a text-to-image diffusion model that takes as input a short text description of an image and then uses AI techniques to generate an image in response to the prompt.

23. In May 2022, Google announced Imagen in a paper called “Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding.”¹ In the paper, Google admits that it trained Imagen on “the publicly available Laion [sic] dataset ... with \approx 400M image-text pairs.”²

24. Initially, Google did not release Imagen to the public. Google explained its reasoning on the website for Imagen: “the data requirements of text-to-image models have led researchers to rely heavily on large, mostly uncurated, web-scraped datasets ... we also utilized LAION-400M dataset which is known to contain a wide range of inappropriate content including pornographic imagery, racist slurs, and harmful social stereotypes ... As such, there is a risk that Imagen has encoded harmful stereotypes and representations, which guides our decision to not release Imagen for public use without further safeguards in place.”³

25. LAION-400M also contains copyrighted works owned by Plaintiffs and the Class, including those in Exhibit B.

¹ Available at <https://arxiv.org/abs/2205.11487>

² *Id.* at 7.

³ See <https://imagen.research.google/>

26. Despite its professed commitment to “not release Imagen for public use without further safeguards,”⁴ Google soon reversed course.

27. In November 2022, Google made Imagen publicly available to a select group of users through its AI Test Kitchen app. According to reporting at the time, Google “announced it will be adding Imagen—in a *very* limited form—to its AI Test Kitchen app as a way to collect early feedback on the technology.”⁵

28. In January 2023, plaintiff Sarah Andersen and two other artists filed the first lawsuit in the U.S. challenging the legality of training text-to-image diffusion models on copyrighted work without consent, credit, or compensation. That case, *Andersen v. Stability AI et al.*, (Case No. 23-cv-00201, N.D. Cal.) challenged two models similar to Imagen—called Stable Diffusion and Midjourney—both of which were also trained on the LAION dataset. (The *Andersen* case is currently proceeding.)

29. In May 2023, Google made Imagen even more widely available through its commercial AI cloud-computing service, called Vertex AI. According to a Google blog post about Vertex AI, Google described it as “Imagen, our text-to-image foundation model, lets organizations generate and customize studio-grade images at scale for any business need.”⁶

30. In October 2023, Google made Imagen even more widely available through a tool called Search Generative Experience. According to reporting at the time, “If you’re opted in to [Search Generative Experience] through Google’s Search Labs program, you can just type your query into the Google search bar. After you do, [Search Generative Experience] can create a few images based on your prompt that you can pick from. The tool is powered by the Imagen family of AI models.”⁷

31. In December 2023, Google released the successor to Imagen, called Imagen 2. Unlike the paper that accompanied the initial version of Imagen, Google’s introduction of Imagen 2 carefully

⁴ *Id.*

⁵ See <https://www.theverge.com/2022/11/2/23434361/google-text-to-image-ai-model-imagen-test-kitchen-app>

⁶ See <https://cloud.google.com/blog/products/ai-machine-learning/google-cloud-launches-new-ai-models-opens-generative-ai-studio>

⁷ See <https://www.theverge.com/2023/10/12/23913337/google-ai-powered-search-sge-images-written-drafts>

omits a detailed description of its training dataset. Google limits itself to vague comments such as “From the outset, we invested in training data safety for Imagen 2, and added technical guardrails to limit problematic outputs like violent, offensive, or sexually explicit content.”⁸

32. On information and belief, Google did not disclose details about the training dataset for Imagen 2 because it was aware of the *Andersen v. Stability AI et al.* case and hoped to avoid being named as a defendant in a lawsuit over the legality of training on mass quantities of copyrighted works without consent, credit, or compensation.

33. On information and belief, Google included LAION-400M in its training dataset for Imagen 2, because a) it had already done so for the first version of Imagen, and b) one of the architects of the LAION image datasets, Romain Beaumont, is a Google employee, who Google hired specifically to exercise influence over the LAION organization and its image datasets.

A KEY SOURCE OF GOOGLE’S TRAINING DATA: LAION

34. LAION (acronym for “Large-Scale Artificial Intelligence Open Network”) is an organization based in Hamburg, Germany. According to its website, LAION is led by Christoph Schuhmann. LAION’s stated goal is “to make large-scale machine learning models, datasets and related code available to the general public.”⁹ All of LAION’s projects are made available for free.

35. Since 2021, a key member of LAION’s team has been Romain Beaumont, who describes himself on the LAION website as an “open source contributor ... I like to apply scale and deep learning to build AI apps and models.”¹⁰

36. LAION’s most well-known projects are the datasets of training images it has released for training machine-learning models, which are now widely used in the AI industry.

37. In August 2021, LAION released LAION-400M, a dataset of 400 million training images assembled from images accessible on the public internet. At the time, LAION-400M was the largest freely available dataset of its kind. Until December 2023, LAION distributed the LAION-400M

⁸ See <https://deepmind.google/technologies/imagen-2/>

⁹ <https://laion.ai/about/>

¹⁰ See <https://laion.ai/team/>

dataset to the public through its own website and elsewhere. (In December 2023, due to the discovery of child sexual-abuse material (“CSAM”) in the LAION datasets, the LAION organization retracted these datasets—including LAION-400M—from the public internet.)

38. Also in August 2021, Romain Beaumont created an online tool called Clip Retrieval that acted as a search interface to LAION to check whether certain artists or artworks were included in the LAION-400M dataset.¹¹ Beaumont’s tool was popular. It was online until December 2023. (In December 2023, it was disabled due to the aforementioned issues with CSAM in the LAION datasets.)

39. In November 2021, Romain Beaumont was a primary author of the paper that introduced the LAION-400M dataset, titled “LAION-400M: Open Dataset of CLIP-Filtered 400 Million Image-Text Pairs,” released in November 2021 (hereafter, the “Beaumont-LAION Paper”).¹²

40. When one downloads the LAION-400M dataset, one gets a list of metadata records, one for each training image. Each record includes the URL of the image, the image caption, a measurement of the similarity of the caption and image, a NSFW flag (indicating the probability the image contains so-called “not safe for work” content), and the width and height of the image.

41. The actual images referenced in the LAION-400M dataset records are not included with the dataset. Anyone who wishes to use LAION-400M for training their own machine-learning model must first acquire copies of the actual images from their URLs. To facilitate the copying of these images, Romain Beaumont created a software tool called `img2dataset` that takes the LAION-400M metadata records as input and makes copies of the referenced images from the URLs in each metadata record, thereby creating local copies. The `img2dataset` tool is distributed from a page Beaumont controls on GitHub.¹³ LAION promotes the `img2dataset` tool in its documentation for LAION-400M. (“This metadata dataset purpose is to download the images for the whole dataset or a subset of it by supplying it to the very efficient `img2dataset` tool.”¹⁴)

¹¹ See <https://rom1504.github.io/clip-retrieval>

¹² <https://arxiv.org/abs/2111.02114>

¹³ <https://github.com/rom1504/img2dataset>

¹⁴ See <https://laion.ai/blog/laion-400-open-dataset/>

42. Training a model with the LAION-400M dataset cannot begin without first using `img2dataset` or another similar tool to download the images in the dataset. Thus, because Google has trained Imagen on LAION-400M, Google has necessarily made one or more copies of images belonging to Plaintiffs as shown in Exhibit B, either by using Romain Beaumont's `img2dataset` tool or another. Plaintiffs never authorized any of these LAION dataset users to copy their images or use them for training any models.

43. One of the entities that has made unauthorized copies of the LAION-400M training images is LAION itself. According to the Beaumont-LAION Paper, LAION made the dataset by starting with Common Crawl metadata records. Common Crawl is a corpus of 250 billion web pages copied from the public web, including assets like Plaintiffs' images (<https://commoncrawl.org/>). The metadata records contain web URLs. According to the Beaumont-LAION Paper, LAION created training images by first "pars[ing] through [the metadata records] from Common Crawl and pars[ing] out all HTML IMG tags containing an alt-text attribute [that is, a text caption]." Then, LAION "download[ed] the raw images from the parsed URLs". Beaumont-LAION Paper at 3.

44. Sometime after the release of LAION-400M in August 2021, a company called Stability AI funded LAION's creation of a similar dataset, but much larger. In March 2022, Stability AI CEO Mostaque called himself "the biggest backer of LAION."¹⁵

45. But Google wasn't far behind. In March 2022, Google hired Romain Beaumont as a full-time software engineer, a position he has held since. On information and belief, Google hired Beaumont primarily to influence the creation of future LAION image datasets, based on a) Beaumont's key role creating LAION-400M—which Google used to train Imagen; b) Beaumont's control of the `img2dataset` tool that was essential to using the LAION-400M dataset, and c) Beaumont's control of the Clip Retrieval website that was essential to searching the LAION-400M dataset.

46. Later in March 2022, LAION released LAION-5B, a dataset of 5.85 billion training images—more than 14 times bigger than LAION-400M. The author of the LAION blog post announcing LAION-5B was Romain Beaumont.¹⁶

¹⁵ <https://discord.com/channels/662267976984297473/938713143759216720/954674533942591510>

¹⁶ See <https://laion.ai/blog/laion-5b/>

47. In August 2022, Romain Beaumont created a specialized AI model to rate the aesthetic quality of an image, and used this model to create subsets of the LAION-5B training images filtered by aesthetic quality, which Beaumont called LAION-Aesthetics. In its introduction of Imagen 2 in December 2023, Google said “We trained a specialized image aesthetics model based on human preferences for qualities like good lighting, framing, exposure, sharpness, and more. Each image was given an aesthetics score which helped condition Imagen 2 to give more weight to images in its training dataset that align with qualities humans prefer.”¹⁷ On information and belief, Beaumont’s work on LAION-Aesthetics formed the basis of Imagen 2’s “aesthetics model”, since at the time Beaumont was both a contributor to LAION and a full-time employee of Google.

48. In October 2022, Romain Beaumont was a primary author of the paper about LAION-5B, called “LAION-5B: An open large-scale dataset for training next generation image-text models.” (hereafter, the “Beaumont-LAION-5B Paper”). According to the Beaumont-LAION-5B Paper, LAION-400M is a subset of LAION-5B, meaning every image in LAION-400M is also in LAION-5B.

49. Just like the LAION-400M dataset, the actual images referenced in the LAION-5B dataset records are not included with the dataset. Anyone who wishes to use LAION-5B for training their own machine-learning model must first acquire copies of the actual images from their URLs. As mentioned above, to facilitate the copying of these images, Romain Beaumont created a software tool called `img2dataset` that takes the LAION-5B metadata records as input and makes copies of the referenced images from the URLs in each metadata record, thereby creating local copies. The `img2dataset` tool is distributed from a page Beaumont controls on GitHub.¹⁸

¹⁷ See <https://deepmind.google/technologies/imagen-2/>

¹⁸ <https://github.com/rom1504/img2dataset>

COUNT 1

DIRECT COPYRIGHT INFRINGEMENT (17 U.S.C. § 501)

AGAINST GOOGLE

50. The preceding factual allegations are incorporated by reference.

51. As the owners of the registered copyrights in the works in Exhibit B, Plaintiffs hold the exclusive rights to those works under the U.S. Copyright Act (17 U.S.C. § 106).

52. Plaintiffs never authorized Google to use their copyrighted work in any way. Nevertheless, Google repeatedly violated Plaintiffs' exclusive rights under § 106 and continues to do so today. Plaintiffs and the Class members never authorized Google to make copies of their works, make derivative works, publicly display copies (or derivative works), or distribute copies (or derivative works).

53. On information and belief, Google has used Plaintiffs' training images to train other versions of Imagen, including Imagen 2, and so-called "multimodal" models that are trained on training images as well as text, such as Google Gemini. Collectively, Imagen and other models that Google trained on LAION-400M are called the **Google-LAION Models**.

54. The LAION-400M dataset contains only URLs of training images, not the actual training images. Therefore, anyone who wishes to use LAION-400M for training their own machine-learning model must first acquire copies of the actual training images from their URLs. Consistent with this, in preparation for training the Google-LAION Models, Google made one or more copies of the LAION-400M training images, including the Plaintiff works in Exhibit B, so they could be fed to the Google-LAION Models as training data. The copies made of each copyrighted work were substantially similar to that copyrighted work.

55. During the training of the Google-LAION Models, Google made a series of intermediate copies of the LAION-400M training images, including the Plaintiff works in Exhibit B. The intermediate copies of each copyrighted work that Google made during training of the Google-LAION Models were substantially similar to that copyrighted work.

56. Plaintiffs have been injured by Google's acts of direct copyright infringement. Plaintiffs are entitled to statutory damages, actual damages, restitution of profits, and other remedies provided by law.

COUNT 2

VICARIOUS COPYRIGHT INFRINGEMENT

AGAINST ALPHABET

57. The preceding factual allegations are incorporated by reference.

58. Alphabet was the corporate parent of Google during its training of the Google-LAION Models and remains its corporate parent.

59. As the corporate parent of Google, Alphabet benefitted financially from the infringing activity of Google when it trained the Google-LAION Models on Plaintiffs' works, and continues to benefit financially from the deployment of the Google-LAION Models.

60. As the corporate parent of Google, Alphabet had the right and ability to supervise the infringing activity of Google when it trained the Google-LAION Models on Plaintiffs' works. Alphabet failed to exercise that right and ability.

61. Plaintiffs have been injured by Alphabet's acts of vicarious copyright infringement. Plaintiffs are entitled to statutory damages, actual damages, restitution of profits, and other remedies provided by law.

CLASS ALLEGATIONS

62. The "**Class Period**" as defined in this Complaint begins on at least April 26, 2021 and runs through the present. Because Plaintiffs do not yet know when the unlawful conduct alleged herein began, but believe, on information and belief, that the conduct likely began earlier than the date listed above, Plaintiffs reserve the right to amend the Class Period to comport with the facts and evidence uncovered during further investigation or through discovery.

63. **Class definition.** Plaintiffs bring this action for damages and injunctive relief as a class action under Federal Rules of Civil Procedure 23(a), 23(b)(2), and 23(b)(3), on behalf of the following Class:

All persons or entities domiciled in the United States that own a United States copyright in any work that Google used as a training image for the Google-LAION Models during the Class Period.

64. This Class definition excludes:

- a. Defendants named herein;
- b. any of the Defendants' co-conspirators;
- c. any of Defendants' parent companies, subsidiaries, and affiliates;
- d. any of Defendants' officers, directors, management, employees, subsidiaries, affiliates, or agents;
- e. all governmental entities; and
- f. the judges and chambers staff in this case, as well as any members of their immediate families.

65. **Numerosity.** Plaintiffs do not know the exact number of members in the Class. This information is in the exclusive control of Defendant. On information and belief, there are at least thousands of members in the Class geographically dispersed throughout the United States. Therefore, joinder of all members of the Class in the prosecution of this action is impracticable.

66. **Typicality.** Plaintiffs' claims are typical of the claims of other members of the Class because Plaintiffs and all members of the Class were damaged by the same wrongful conduct of Defendant as alleged herein, and the relief sought herein is common to all members of the Class.

67. **Adequacy.** Plaintiffs will fairly and adequately represent the interests of the members of the Class because the Plaintiffs have experienced the same harms as the members of the Class and have no conflicts with any other members of the Class. Furthermore, Plaintiffs have retained sophisticated and competent counsel who are experienced in prosecuting federal and state class actions, as well as other complex litigation.

68. **Commonality and predominance.** Numerous questions of law or fact common to each Class member arise from Defendants' conduct and predominate over any questions affecting the members of the Class individually:

- a. Whether Defendants violated the copyrights of Plaintiffs and the Class when they obtained copies of Plaintiffs' copyrighted images and used them to train the Google-LAION Models.
- b. Whether any affirmative defense excuses Defendants' conduct.
- c. Whether any statutes of limitation constrain the potential for recovery for Plaintiffs and the Class.

69. **Other class considerations.** Defendants have acted on grounds generally applicable to the Class. This class action is superior to alternatives, if any, for the fair and efficient adjudication of this controversy. Prosecuting the claims pleaded herein as a class action will eliminate the possibility of repetitive litigation. There will be no material difficulty in the management of this action as a class action.

70. The prosecution of separate actions by individual Class members would create the risk of inconsistent or varying adjudications, establishing incompatible standards of conduct for Defendants.

DEMAND FOR JUDGMENT

WHEREFORE, Plaintiffs request that the Court enter judgment on their behalf and on behalf of the Class defined herein, by ordering:

- a) This action may proceed as a class action, with Plaintiffs serving as Class Representatives, and with Plaintiffs' counsel as Class Counsel.
- b) Judgment in favor of Plaintiffs and the Class and against Defendant.
- c) An award of statutory and other damages under 17 U.S.C. § 504 for violations of the copyrights of Plaintiffs and the Class by Defendant.
- d) Destruction or other reasonable disposition of all copies Defendants made or used in violation of the exclusive rights of Plaintiffs and the Class, under 17 U.S.C. § 503(b).

- e) Pre- and post-judgment interest on the damages awarded to Plaintiffs and the Class, and that such interest be awarded at the highest legal rate from and after the date this class action complaint is first served on Defendant.
- f) Defendants are to be jointly and severally responsible financially for the costs and expenses of a Court approved notice program through post and media designed to give immediate notification to the Class.
- g) Further relief for Plaintiffs and the Class as may be just and proper.

JURY TRIAL DEMANDED

Under Federal Rule of Civil Procedure 38(b), Plaintiffs demand a trial by jury of all the claims asserted in this Complaint so triable.

Dated: April 26, 2024

By: /s/ Joseph R. Saveri
Joseph R. Saveri

Joseph R. Saveri (State Bar No. 130064)
Cadio Zirpoli (State Bar No. 179108)
Christopher K. L. Young (State Bar No. 318371)
Elissa Buchanan (State Bar No. 249996)
JOSEPH SAVERI LAW FIRM, LLP
601 California Street, Suite 1505
San Francisco, CA 94108
Telephone: (415) 500-6800
Facsimile: (415) 395-9940
Email: jsaveri@saverilawfirm.com
czirpoli@saverilawfirm.com
cyoung@saverilawfirm.com
eabuchanan@saverilawfirm.com

Matthew Butterick (State Bar No. 250953)
1920 Hillhurst Avenue, #406
Los Angeles, CA 90027
Telephone: (323) 968-2632
Facsimile: (415) 395-9940
Email: mb@buttericklaw.com

Brian D. Clark (*pro hac vice* pending)
Laura M. Matson (*pro hac vice* pending)
Arielle S. Wagner (*pro hac vice* pending)
Eura Chang (*pro hac vice* pending)
LOCKRIDGE GRINDAL NAUEN PLLP
100 Washington Avenue South, Suite 2200
Minneapolis, MN 55401
Telephone: (612) 339-6900
Facsimile: (612) 339-0981
Email: bdclark@locklaw.com
lmmatson@locklaw.com
aswagner@locklaw.com
echang@locklaw.com

*Counsel for Individual and Representative
Plaintiffs and the Proposed Class*

Exhibit A: Plaintiff Copyright Registrations

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Text

Registration Number / Date:
TX0008629478 / 2018-05-15

Application Title: ALL SUMMER LONG.

Title: ALL SUMMER LONG.

Description: Book, 170 p.

Copyright Claimant:
Hope Larson.

Date of Creation: 2018

Date of Publication:
2018-04-12

Nation of First Publication:
United States

Authorship on Application:
Hope Larson; Domicile: United States. Authorship: text,
artwork.

Rights and Permissions:
Macmillan Children's Publishing Group, Attn: Permissions
Manager, 175 Fifth Avenue, New York, NY, 10010, United
States

Names: Larson, Hope

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Text

Registration Number / Date:
TX0006879805 / 2008-09-11

Application Title: Chiggers.

Title: Chiggers.

Description: Book, 170p.

Copyright Claimant:
Hope Larson

Date of Creation: 2008

Date of Publication:
2008-06-17

Nation of First Publication:
United States

Authorship on Application:
Hope Larson; Citizenship: United States. Authorship: text,
artwork.

Rights and Permissions:
To whom it may concern, Simon & Schuster, Simon & Schuster,
Inc.-17th Floor, 1230 Avenue of the Americas, New York,
NY, 10020

ISBN: 1416935843

Names: Larson, Hope

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Visual Material

Registration Number / Date:
VA0002097201 / 2017-07-13

Application Title: Chester 5000 XYV, Book 2: Isabelle & George HC.

Title: Chester 5000 XYV, Book 2: Isabelle & George HC.

Description: Book, 180 p.

Copyright Claimant:
Jessica Fink.

Date of Creation: 2016

Date of Publication:
2016-05-01

Nation of First Publication:
United States

Authorship on Application:
Jessica Fink; Domicile: United States. Authorship: 2-D
artwork, Text.

Names: Fink, Jessica
Fink, Jess

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Visual Material

Registration Number / Date:
VA0001911930 / 2014-04-08

Application Title: Anouk, et al.

Title: Anouk.

Appears in: Fashion Gone Rogue, Nov 12, 2013

Description: Electronic file (eService)

Copyright Claimant:
Jingna Zhang, 1988- .

Date of Creation: 2013

Date of Publication:
2013-02-17

Nation of First Publication:
United States

Authorship on Application:
Jingna Zhang, 1988- ; Citizenship: Singapore. Authorship:
photograph(s)

Rights and Permissions:
Jingna Zhang, info@zhangjingna.com

Names: Zhang, Jingna, 1988-

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Visual Material

Registration Number / Date:
VA0001917341 / 2014-06-09

Application Title: Motherland Chronicles.

Title: Motherland Chronicles.

Description: Electronic file (eService)

Copyright Claimant:
Jingna Zhang, 1988- .

Date of Creation: 2013

Date of Publication:
2013-02-04

Nation of First Publication:
United States

Authorship on Application:
Jingna Zhang, 1988- ; Citizenship: Singapore. Authorship:
photograph(s), 2-D artwork.

Rights and Permissions:
Jingna Zhang, info@zhangjingna.com

Names: Zhang, Jingna, 1988-

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Text

Registration Number / Date:
TX0009082399 / 2021-01-22

Application Title: Adulthood is a Myth: A "Sarah's Scribbles" 2021 Wall
Calendar.

Title: Adulthood is a Myth: A "Sarah's Scribbles" 2021 Wall
Calendar.

Description: Book.

Copyright Claimant:
Sarah Andersen.

Date of Creation: 2020

Date of Publication:
2020-06-30

Nation of First Publication:
United States

Authorship on Application:
Sarah Andersen; Citizenship: United States. Authorship:
artwork, compilation of previously published art with
accompanying text; some new art.

Pre-existing Material:
previously published art with accompanying text.

Basis of Claim: artwork, compilation of previously published art with
accompanying text; some new art.

ISBN: 9781524857677

Names: Andersen, Sarah

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Text

Registration Number / Date:
TX0008207926 / 2015-12-10

Application Title: Adulthood Is a Myth: A Sarah's Scribbles Collection.

Title: Adulthood Is a Myth: A Sarah's Scribbles Collection.

Description: Book.

Copyright Claimant:
Sarah Andersen.

Date of Creation: 2015

Date of Publication:
2015-12-08

Nation of First Publication:
United States

Authorship on Application:
Sarah Andersen; Citizenship: United States. Authorship:
text, artwork.

Pre-existing Material:
Some cartoons previously appeared on author's website.

Basis of Claim: text, artwork.

ISBN: 9781449474195

Names: Andersen, Sarah

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Text

Registration Number / Date:
TX0008493928 / 2017-01-09

Application Title: Big Mushy Happy Lump: A Sarah's Scribbles Collection .

Title: Big Mushy Happy Lump: A Sarah's Scribbles Collection .

Description: Book, 125 p.

Copyright Claimant:
Sarah Andersen.

Date of Creation: 2016

Date of Publication:
2016-12-12

Nation of First Publication:
United States

Authorship on Application:
Sarah Andersen; Citizenship: United States. Authorship:
text, artwork.

Copyright Note: Basis for Registration: Collective work

ISBN: 9781449479619

Names: Andersen, Sarah

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Text

Registration Number / Date:
TX0008564573 / 2018-02-12

Application Title: Herding Cats: A Sarah's Scribbles Collection.

Title: Herding Cats: A Sarah's Scribbles Collection.

Description: Book, 108 p.

Copyright Claimant:
Sarah Andersen.

Date of Creation: 2018

Date of Publication:
2018-01-16

Nation of First Publication:
United States

Authorship on Application:
Sarah Andersen; Citizenship: United States. Authorship:
text, artwork, Curated and themed collection of cartoons
previously published online.

Pre-existing Material:
Some cartoons previously published online.

Basis of Claim: text, artwork, Curated and themed collection of cartoons
previously published online.

ISBN: 9781449489786

Names: Andersen, Sarah

Exhibit A: Plaintiff Copyright Registrations

Type of Work: Text

Registration Number / Date:
TX0009153980 / 2022-02-16

Application Title: Oddball: A Sarah's Scribbles Collection.

Title: Oddball: A Sarah's Scribbles Collection.

Description: Book, 105 p.

Copyright Claimant:
Sarah Andersen.

Date of Creation: 2021

Date of Publication:
2021-12-07

Nation of First Publication:
United States

Authorship on Application:
Sarah Andersen; Citizenship: United States. Authorship:
text, artwork.

Copyright Note: Basis for Registration: Collective work

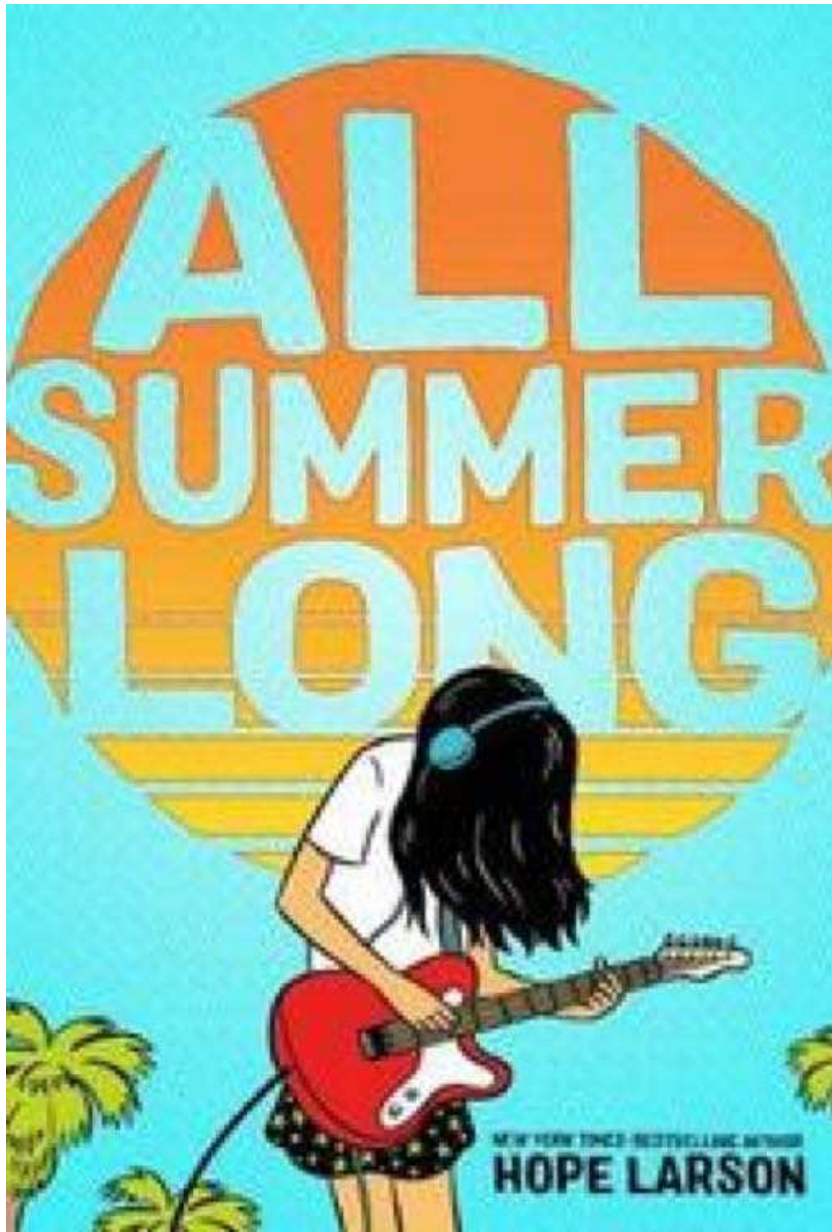
ISBN: 9781449489793

Names: Andersen, Sarah

Exhibit B: Plaintiff Images in LAION-400M

Exhibit B: Plaintiff Images in LAION-400M

Hope Larson: training image 1



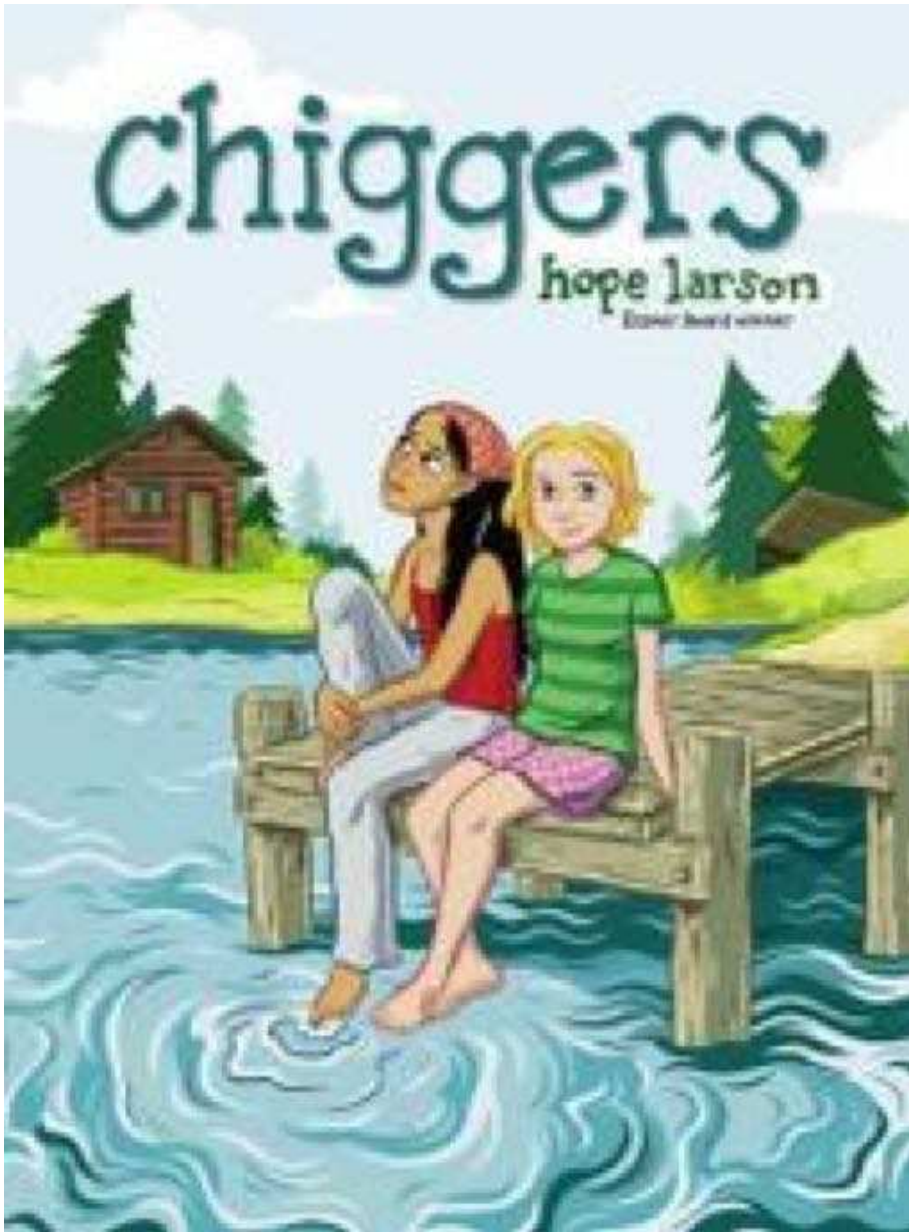
copyright registration: All Summer Long (TX0008629478)

LAION-400M caption: "All Summer Long by Hope Larson"

LAION-400M url: <https://d3fa68hw0m2vcc.cloudfront.net/471/196258371.jpeg>

Exhibit B: Plaintiff Images in LAION-400M

Hope Larson: training image 2



copyright registration: Chiggers (TX0006879805)

LAION-400M caption: "CHIGGERS by Hope Larson"

LAION-400M url: <https://d1ldy8a769gy68.cloudfront.net/180/978/141/693/584/1/9781416935841.jpg>

Exhibit B: Plaintiff Images in LAION-400M

Jessica Fink: training image 1



copyright registration: Chester 5000: Isabelle & George

LAION-400M caption: "Chester 5000 (Book 2): Isabelle & George"

LAION-400M url: http://cdn.topshelfcomix.com/catalog/covers/chester_5000_book_2_cover_sm.jpg

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 1



copyright registration: Motherland Chronicles (VA0001917341)

LAION-400M caption: "aurorae: Motherland Chronicles #18 - Julia by `zemotion on deviantART"

LAION-400M url: <https://i.pinimg.com/736x/8b/6c/23/8b6c239d46e3a49d2ead483b9949b8c0--character-ideas-character-design.jpg>

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 2



copyright registration: Motherland Chronicles (VA0001917341)

LAION-400M caption: "Motherland Chronicles - Self Portrait in Water , 2013"

LAION-400M url: [https://images.squarespace-cdn.com/content/53adb9cbe4b05e0d792c6c88/1405552099182-](https://images.squarespace-cdn.com/content/53adb9cbe4b05e0d792c6c88/1405552099182-GVWPC3BG8KXVER3AUGVV/Motherland-Chronicles-7---Self-Portrait-in-Water-Zhang-Jingna-zemotion.jpg?content-type=image%2Fjpeg)

[GVWPC3BG8KXVER3AUGVV/Motherland-Chronicles-7---Self-Portrait-in-Water-Zhang-Jingna-zemotion.jpg?content-type=image%2Fjpeg](https://images.squarespace-cdn.com/content/53adb9cbe4b05e0d792c6c88/1405552099182-GVWPC3BG8KXVER3AUGVV/Motherland-Chronicles-7---Self-Portrait-in-Water-Zhang-Jingna-zemotion.jpg?content-type=image%2Fjpeg)

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 3



copyright registration: Motherland Chronicles (VA0001917341)

LAION-400M caption: "Motherland Chronicles 2 - Winterland Fairytales (by zemotion) "

LAION-400M url: https://64.media.tumblr.com/756ad0716e5385edcd9a7e88329df781/tumblr_miauw7P6Yt1qafz6uo1_500.jpg

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 4



copyright registration: Motherland Chronicles (VA0001917341)

LAION-400M caption: "Motherland Chronicles - Book of Roses , 2013 "

LAION-400M url: [https://images.squarespace-cdn.com/content/53adb9cbe4b05e0d792c6c88](https://images.squarespace-cdn.com/content/53adb9cbe4b05e0d792c6c88/1405551976391-0QPVOVJGZNORCB97DEUE/Motherland-Chronicles-31---Book-of-Roses-Zhang-Jingna-zemotion.jpg?content-type=image%2Fjpeg)

[/1405551976391-0QPVOVJGZNORCB97DEUE/Motherland-Chronicles-31---Book-of-Roses-Zhang-Jingna-zemotion.jpg?content-type=image%2Fjpeg](https://images.squarespace-cdn.com/content/53adb9cbe4b05e0d792c6c88/1405551976391-0QPVOVJGZNORCB97DEUE/Motherland-Chronicles-31---Book-of-Roses-Zhang-Jingna-zemotion.jpg?content-type=image%2Fjpeg)

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 5



copyright registration: Motherland Chronicles (VA0001917341)

LAION-400M caption: "Motherland-Chronicles #24 - Dark Alodia "

LAION-400M url: https://images-wixmp-ed30a86b8c4ca887773594c2.wixmp.com/f/bbf9e125-bd9e-4a70-badc-1bd5affeb49f/d6dpzv3-65e2c333-9b9c-43da-b6d0-3bbe25fdd858.jpg/v1/fill/w_221,h_350,q_70,strp/motherland_chronicles__24__dark_alodia_by_zemotion_d6dpzv3-350t.jpg?token=eyJ0eXAiOiJKV1QiLCJhbGciOiJIUzI1NiJ9.eyJzdWIiOiJ1cm46YXBwOjdlMGQxODg5ODIyNjQzNzNhNWYwZDQxNWVhMGQyNmUwIiwiaXNzIjoidXJuOmFwcDo3ZTBkMTg4OTgyMjY0MzczYTVmMGQ0MTVIYTBlbmJZIMCislm9iail6W1t7ImhlaWdodCI6Ijw9Nzk1IiwicGF0aCI6IiwvZlwwYmJmOWUxMjUtYmQ5ZS00YTcwLWJhZGMtMWJkNWVmZmViNDlmXC9kNmRwenYzLTlY1ZTJjMzMzLTliOWMtNDNkYS1iNmQwLTNiYmUyNWZkZDg1OC5qcGciLCJ3aWR0aCI6Ijw9NTAyIn1dXSwiYXVkljpbInVybjpzZXJ2aWNlOmItYWdlLm9wZXJhdGlbnMiX30.B3RPZhY7HYihH0yxgTkcqmbouHq5QNlepHxW03bCy1s

bcfg-34be19b0a39e.jpg?token=eyJ0eXAiOiJKV1QiLCJhbGciOiJIUzI1NiJ9.eyJpc3MiOiJ1cm46YXBwOjldMGQxODg5ODIyNjQzNzNhNWYwZDQxNWVhMGQyNmUwIiwic3ViIjoiaXJuOmFwcDo3ZTBkMTg4OTgyMjY0MzczYTVmMGQ0MTVlYTbkMjZlMCIsmF1ZCI6WYJ1cm46c2VydmljZTpmWwXlMmRvd25sb2Fkl0slm9iaill6W1t7lnBhdGgiOiIlvZi9iYmY5ZTEyNS1iZDlLTRhZnAtYmFkYyOxYmQ1YWZmZWl0OWYvZDZjb3dzYi0xYjMxZWY0My1mMGQ0LTriMmEtYmNmNi0zNGJlMTliMGZlOWUuanBnIn1dXX0.5kPKgrrcV8_KFb1iK6wpRimWvGgPLLfxY2FJhWfqgR

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 7



copyright registration: Motherland Chronicles (VA0001917341)

LAION-400M caption: "Motherland-Chronicles-33---Underwater-Ascend-Zhang-Jingna-zemotion.jpg"

LAION-400M url: <https://images.squarespace-cdn.com/content/55db733fe4b0725b23e574a4/1529261106379-79BSSEUNOWD8JD2W2EDF/Motherland-Chronicles-33---Underwater-Ascend-Zhang-Jingna-zemotion.jpg?content-type=image%2Fjpeg>

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 8



copyright registration: Anouk (VA0001911930)

LAION-400M caption: "Fashion Gone Rogue: Anouk"

LAION-400M url: <https://mir-s3-cdn-cf.behance.net/projects/source/12146521.5482638fe6b81.jpg>

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 9



copyright registration: Anouk (VA0001911930)

LAION-400M caption: "anouk van kleef6 Anouk van Kleef by Zhang Jingna for Fashion Gone Rogue "

LAION-400M url: <https://www.fashiongonerogue.com/wp-content/uploads/2013/11/anouk-van-kleef6.jpg>

Exhibit B: Plaintiff Images in LAION-400M

Jingna Zhang: training image 10



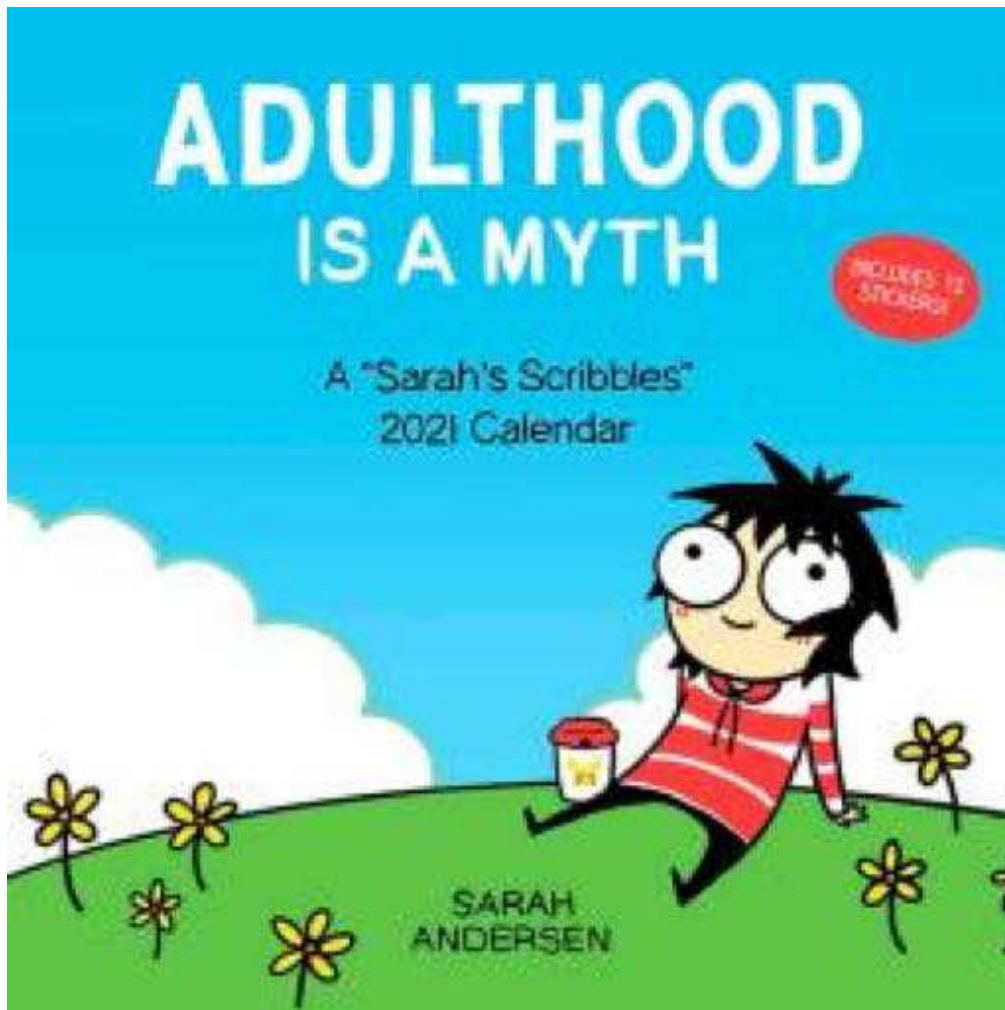
copyright registration: Anouk (VA0001911930)

LAION-400M caption: "anouk van kleeft4 Anouk van Kleef by Zhang Jingna for Fashion Gone Rogue"

LAION-400M url: <https://www.fashiongonerogue.com/wp-content/uploads/2013/11/anouk-van-kleeft4.jpg>

Exhibit B: Plaintiff Images in LAION-400M

Sarah Andersen: training image 1



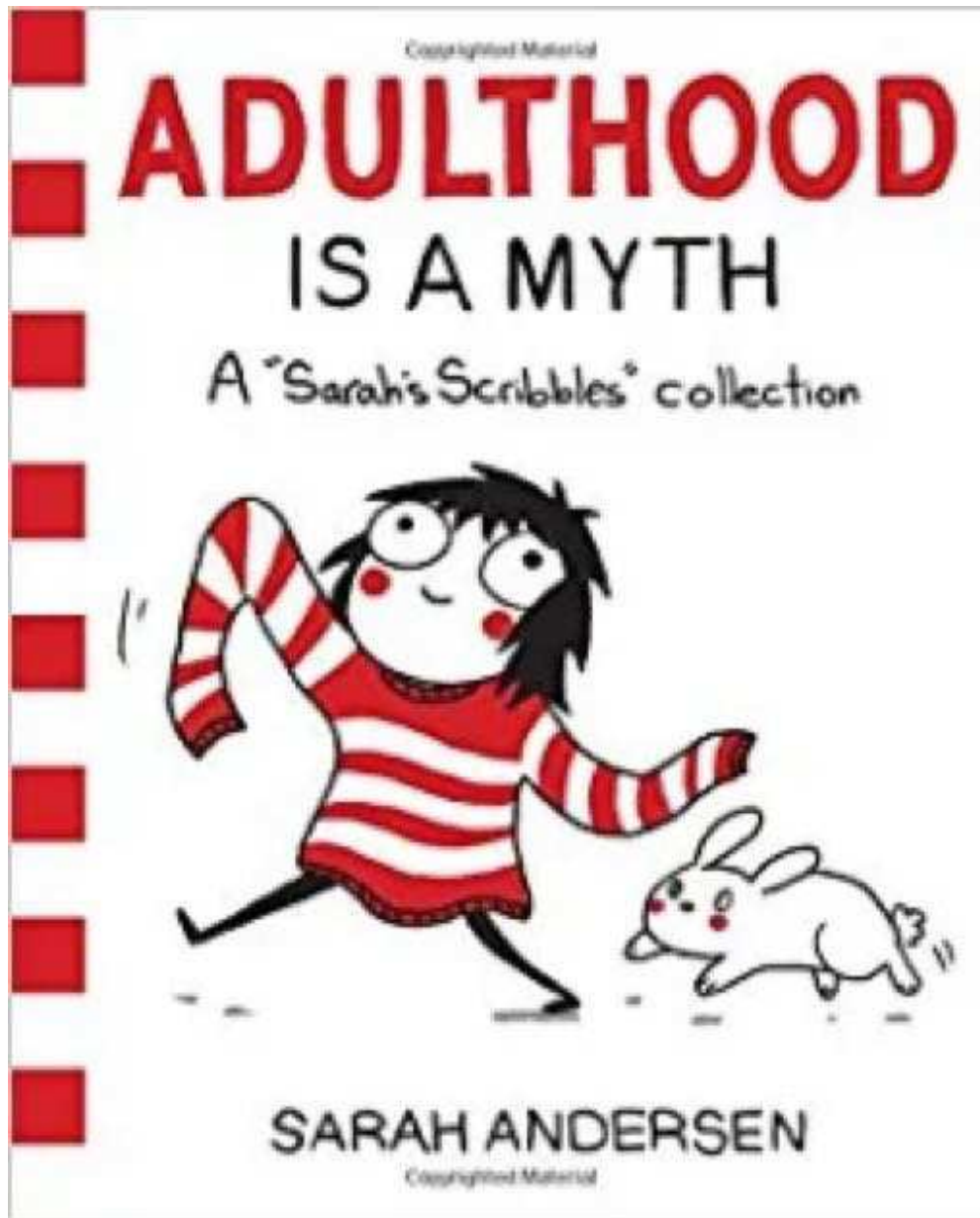
copyright registration: Adulthood is a Myth: A "Sarah's Scribbles" 2021 Wall Calendar (TX0009082399)

LAION-400M caption: "Adulthood is a Myth: A Sarah's Scribbles 2021 Wall Calendar"

LAION-400M url: <https://www.sfbok.se/sites/default/files/styles/teaser/sfbok/sfbokbilder/207/207606.jpg?bust=1582127788&itok=Ej3ZNa8z>

Exhibit B: Plaintiff Images in LAION-400M

Sarah Andersen: training image 2



copyright registration: Adulthood Is a Myth: A Sarah's Scribbles Collection (TX0008207926)

LAION-400M caption: "adulthood is a myth sarah andersen 2016 andrews mcmeel publishing"

LAION-400M url: <https://omniramblings.files.wordpress.com/2017/09/adulthood-is-a-myth.jpg?w=244&>

Exhibit B: Plaintiff Images in LAION-400M

Sarah Andersen: training image 3



copyright registration: Big Mushy Happy Lump: A Sarah's Scribbles Collection (TX0008493928)

LAION-400M caption: "I have so much work to do. | Sarah's Scribbles By sarahseeandersen
[Comic - Drawing - Illustration - Funny]"

LAION-400M url: <https://i.pinimg.com/236x/65/fd/f4/65fdf4c41d21a5ca3d33379311455805--sarah-andersen-comics-sarah-see-andersen.jpg>

Exhibit B: Plaintiff Images in LAION-400M

Sarah Andersen: training image 4



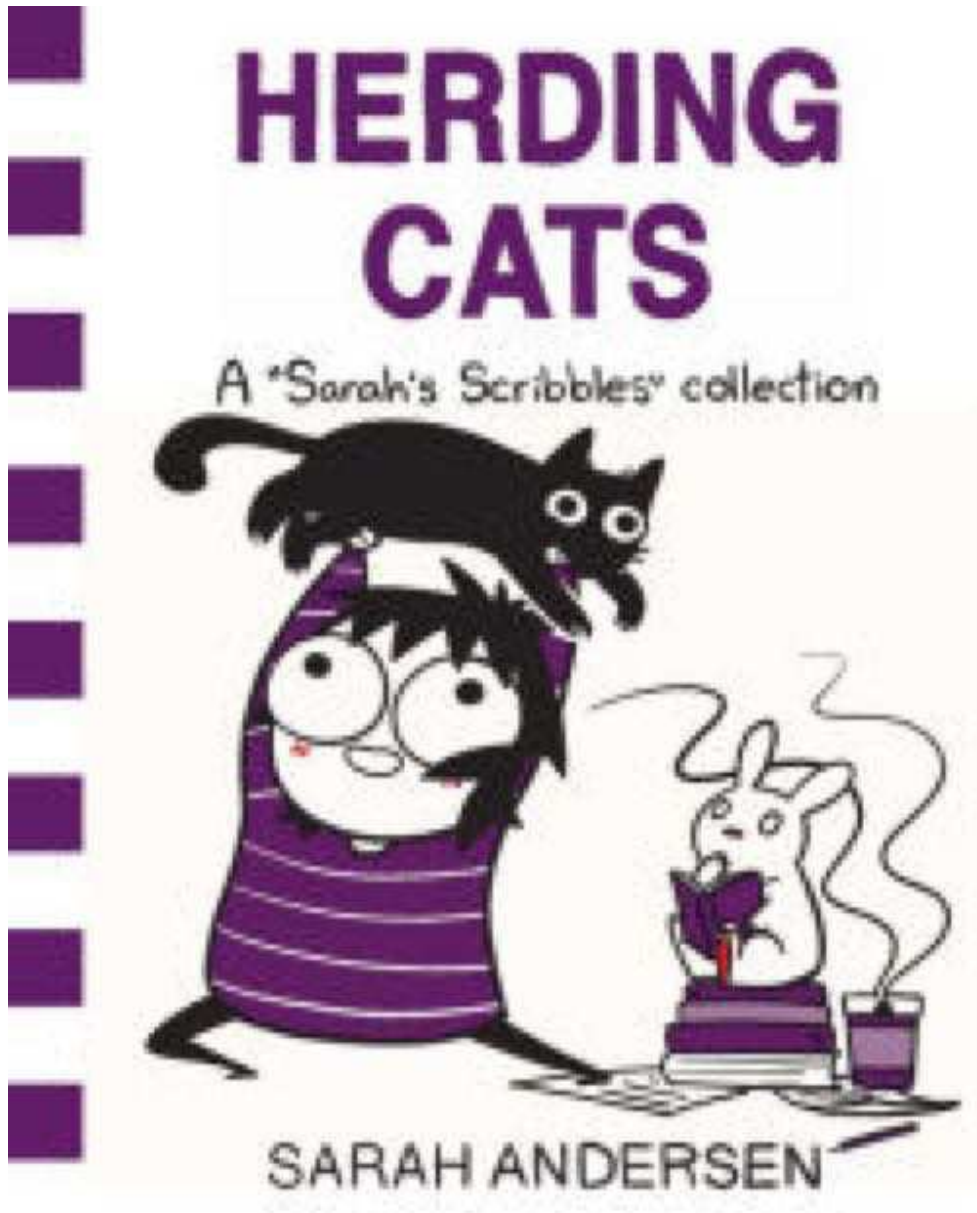
copyright registration: Big Mushy Happy Lump: A Sarah's Scribbles Collection (TX0008493928)

LAION-400M caption: ""Sarah Andersen comic ""How I Spend Money""""

LAION-400M url: <https://img.saplingcdn.com/375/ppds/5496bebe-8fd4-47d0-9538-6063abcbceaf.jpg>

Exhibit B: Plaintiff Images in LAION-400M

Sarah Andersen: training image 5



copyright registration: Herding Cats: A Sarah's Scribbles Collection (TX0008564573)

LAION-400M caption: "Title: Herding Cats: A Sarah's Scribbles Collection, Author: Sarah Andersen"

LAION-400M url: https://prodimage.images-bn.com/pimages/9781449489786_p0_v2_s192x300.jpg

Exhibit B: Plaintiff Images in LAION-400M

Sarah Andersen: training image 6



copyright registration: Oddball: A Sarah's Scribbles Collection (TX0009153980)

LAION-400M caption: "All credit goes to Sarah Andersen - meme"

LAION-400M url: <https://images7.memedroid.com/images/UPLOADED671/5fd3dd71dc22b.jpeg>